

# A Cascading Failure Model by Quantifying Interactions

Junjian Qi and Shengwei Mei

*Department of Electrical Engineering, Tsinghua University, Beijing, China 100084*

Cascading failures triggered by trivial initial events are encountered in many complex systems. It is the interaction and coupling between components of the system that causes cascading failures. We propose a simple model to simulate cascading failure by using the matrix that determines how components interact with each other. A careful comparison is made between the original cascades and the simulated cascades by the proposed model. It is seen that the model can capture general features of the original cascades, suggesting that the interaction matrix can well reflect the relationship between components. An index is also defined to identify important links and the distribution follows an obvious power law. By eliminating a small number of most important links the risk of cascading failures can be significantly mitigated, which is dramatically different from getting rid of the same number of links randomly.

Cascading failures in complex systems are complicated sequences of dependent outages. They can take place in electric power systems [1, 2], the Internet [3, 4], the road system [5, 6], and the social and economic systems [7]. For example, in power grid a line is tripped for some reason, such as error of operators, bad weather, or tree contact, and the power transmitted through this line will be redistributed to other lines, possibly causing other line tripping or even cascading failures.

Several models have been proposed to study the general mechanisms of cascading failures. The topological threshold models [8–10] provide insight into several types of cascading failures. The influence model [11] comprises sites connected by a network structure and each site has a status evolving according to Markov chain. The line interaction graph [12] initiates a novel analysis method for cascading failures in electric power systems by considering the interaction of the transmission lines.

The system-level failure of tightly coupled complex systems is not caused by any specific reason but the property that components of the system are tightly coupled and dependent [13]. Based on this idea we quantify the interaction of the components by an interaction matrix and then propose a cascading failure model by using this matrix. The interaction matrix can be obtained with cascades from simulations or real statistical data. These cascades can be grouped into different generations [14–16] and are called original cascades to distinguish with the simulated cascades in the following discussion. A typical cascade can be:

generation 0	generation 1
25, 40, 74, 102, 155	72, 73, 82

Here the numbers are the serial number of the failed components. This cascade has two generations while others might contain one or several generations.

Topological properties such as small-world [17] and scale-free [18] behavior have been found in complex networks. Power-law behavior is also discovered for weighted

networks [19, 20]. In this paper we discuss the property of the directed weighted interaction network rather than the network directly from the physical system.

Assume there are  $m$  components in the system, we construct matrix  $\mathbf{A} \in \mathbb{Z}^{m \times m}$  whose entry  $a_{ij}$  is the number of times that component  $i$  fails in one generation before the failure of component  $j$  among all original cascades. For each component failure in this generation we find a component failure that most probably causes it. Specifically, the failure of component  $j$  is considered to be caused by the component failure with the greatest  $a_{ij}$  among all component failures in the last generation, thus correcting  $\mathbf{A}$  to be  $\mathbf{A}' \in \mathbb{Z}^{m \times m}$ , whose entry  $a'_{ij}$  is the number of times that the failure of component  $i$  causes the failure of component  $j$ . An example is given in Fig. 1. Then we can get the  $m \times m$  interaction matrix  $\mathbf{B} \in \mathbb{R}^{m \times m}$  whose entry  $b_{ij}$  is the empirical probability that the failure of component  $i$  causes the failure of component  $j$ . From the Bayes' theorem we have  $b_{ij} = a'_{ij}/f_i$ , where  $f_i$  is the number of times that component  $i$  fails. The  $\mathbf{B}$  matrix actually determines how components interacts with each other.

We propose a simple model to simulate cascading failures with  $\mathbf{B}$  matrix. Initially all components are assumed to work well and the cascading failure is triggered by a small fraction of component failures. Since we would like to focus on the interaction of components rather than the triggering events, the component failures in generation 0 of an original cascade are directly considered as generation 0 failures in the simulated cascade. The columns of  $\mathbf{B}$  corresponding to the initial failures are set zero since in our model once a component fails it will remain that way until the end of the simulation. Then the component failures in generation 0 independently generate other component failures. Specifically, if component  $i$  fails in generation 0 it will cause the failure of any other component  $j$  with probability  $b_{ij}$ . Once it causes the failure of a component, the column of  $\mathbf{B}$  corresponding to that component will be set zero. All component failures caused by generation 0 failures comprise generation 1. Generation 1 failures then generate generation 2. This continues

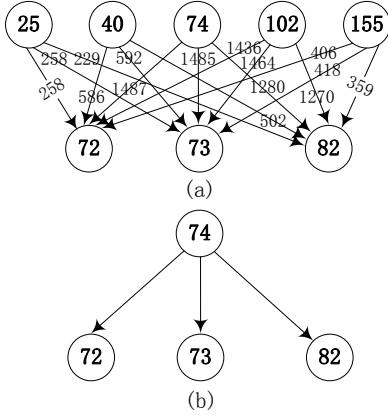


FIG. 1. Illustration of correcting  $\mathbf{A}$  to  $\mathbf{A}'$ . Two consecutive generations of a cascade are given. The last generation comprises the failure of 25, 40, 74, 102, and 155 and in this generation there are three component failures, which are 72, 73, and 82. (a) The numbers on the edges are  $a_{ij}$ , where  $i$  is the source vertex and  $j$  is the destination vertex. (b) When constructing  $\mathbf{A}'$ , for the two consecutive generations 72, 73, and 82 are considered to be caused by 74 since the  $a_{ij}$  starting from 74 are the greatest.

until no failure is caused.

An example for electric power systems is presented to illustrate the proposed model. The original cascades are generated by the Alternating Current ORNL-Pserc-Alaska (AC OPA) simulation [21–23], which is the AC counterpart of the OPA model [24, 25]. We use the form of the AC OPA simulation in which the power system is fixed and does not evolve or upgrade. As other variants of the OPA model, AC OPA can also naturally produce line outages in generations; each iteration of the “main loop” of the simulation produces another generation. We choose transmission lines as components and simulate a total of 5000 cascades on the IEEE 118 bus system [26], which represents a portion of a past American Electric Power Company transmission system.

The complementary cumulative distributions (CCD) of the total number of line outages for original and simulated cascades are shown in Fig. 2. It is seen that the two distributions match well. In our case  $\mathbf{B}$  is a rather sparse matrix, a  $186 \times 186$  matrix with only 202 nonzero elements. Just because of the interaction between components denoted by this sparse matrix, the cascading failure is able to propagate a lot, which is suggested by the dramatic difference between the generation 0 distribution and the total line outage distribution. If the elements of  $\mathbf{B}$  are all zeros and the components do not interact, all cascades will stop immediately after generation 0 failures and the distribution of the total line outages will be the same as generation 0 failures.

The nonzero elements of  $\mathbf{B}$  determine how one component affects another. They are called *links*. A link  $l : i \rightarrow j$  corresponds to the nonzero element  $b_{ij}$  and

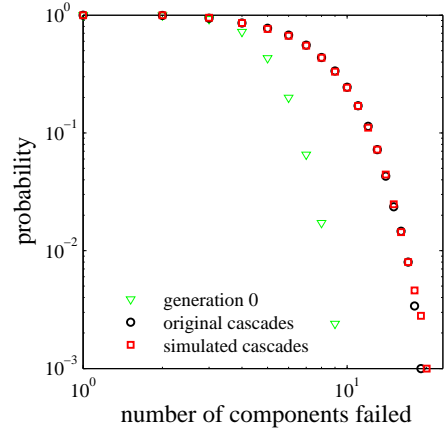


FIG. 2. CCD of the total number of line outages for original cascades (circles) and simulated cascades (rectangles). Downward-pointing triangles denote CCD of the generation 0 failures.

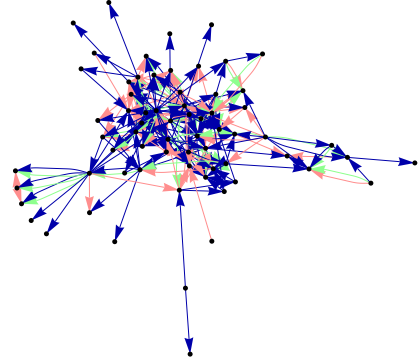


FIG. 3. Links for the original and simulated cascades. Blue links are shared by both original and simulated cascades; red links are only for original cascades; green links are only for simulated cascades; dots denote component failures.

starts from the failure of component  $i$  and ends with the failure of component  $j$ . These links form a directed network, for which the vertices are component failures and the directed links represent that the source vertex causes the destination vertex with probability greater than 0. Fig. 3 shows the network for both original and simulated cascades. The shared links belong to set  $L_1$  and the links only owed by the original and simulated cascades respectively belong to set  $L_2$  and  $L_3$ . In our case 133 links are shared by the original and simulated cascades. 69 links are owned only by the original cascades and 45 only by the simulated cascades.

It seems that the simulated cascades are quite different from the original cascades since there are many different links between them. However, it is not the truth if we take into account the intensities of the links. The links

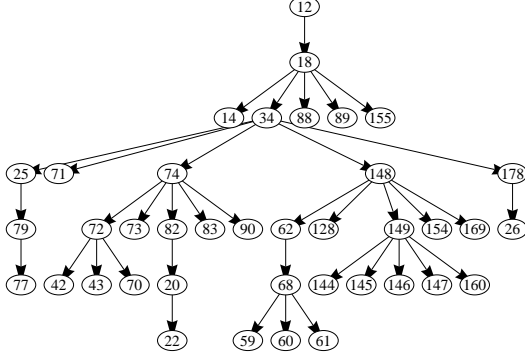


FIG. 4. Diagram showing how  $I_l$  is calculated.

are not equally important and can cause dramatically different consequences. We quantify the importance of a link  $l : i \rightarrow j$  by an index  $I_l$ , which measures the expected component failures that a specific link  $l$  can cause on the condition that the number of times that its source vertex fails is known.

Specifically, assume a component  $i$  fails for  $N_i$  times, the expected failures of component  $j$  will be  $N_j = N_i b_{ij}$ . The expected failures caused by the failure of component  $j$  is  $N_j \sum_{k \in j} b_{jk}$ , where  $k \in j$  denotes the destination vertices starting from  $j$ . We continue to calculate the expected failures until reaching the leaf vertices. All the expected failures are summated to be  $I_l$ . In fact,  $I_l = \sum_{v \in V} N_v$ , where  $V$  is the set of vertices for which there exists a path starting from link  $l$  and  $N_v$  is the expected failures of vertex  $v$ . Fig. 4 lists all vertices that can be influenced by link  $12 \rightarrow 18$ .

By using  $I_l$  as weights of the links we can get a directed weighted network corresponding to the nonzero elements of  $\mathbf{B}$  matrix. Denote the index of link  $l$  for the original and simulated cascades respectively as  $I_l^{\text{ori}}$  and  $I_l^{\text{sim}}$ .  $\sum_{l \in (L_1 \cup L_3)} I_l^{\text{sim}} / \sum_{l \in (L_1 \cup L_2)} I_l^{\text{ori}}$  is 0.9917, which means that the links of the original and simulated cascades have almost the same propagation capacity on the whole. Besides,  $\sum_{l \in L_1} I_l^{\text{ori}} / \sum_{l \in (L_1 \cup L_2)} I_l^{\text{ori}}$  and  $\sum_{l \in L_1} I_l^{\text{sim}} / \sum_{l \in (L_1 \cup L_3)} I_l^{\text{sim}}$  are separately 0.9690 and 0.9620, indicating that the shared links play the major role among all links.

For the shared links  $\sum_{l \in L_1} I_l^{\text{sim}} / \sum_{l \in L_1} I_l^{\text{ori}}$  is 0.9845. This suggests that the overall effects of the shared links of the original and simulated cascades are close to each other. However, it is still possible that the weights of the same link for the original and simulated cascades can be quite different. To show if the same link is close to each other we define a similarity index  $S$  for the shared links.

$$S = \sum_{l \in L_1} \left( \frac{I_l^{\text{sim}} + I_l^{\text{ori}}}{\sum_{l \in L_1} (I_l^{\text{sim}} + I_l^{\text{ori}})} \frac{I_l^{\text{sim}}}{I_l^{\text{ori}}} \right) \quad (1)$$

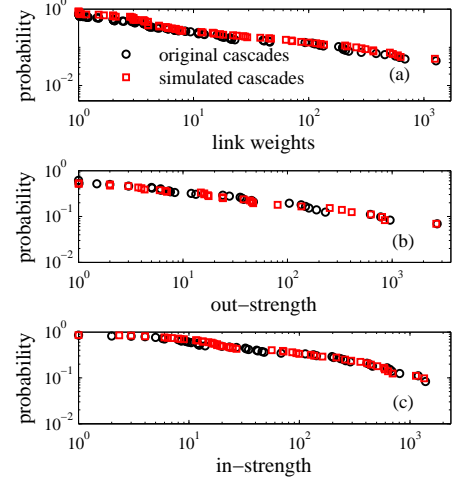


FIG. 5. (a) CCD of the link weights. (b) CCD of the vertex out-strength. (c) CCD of the vertex in-strength.

When the important links of the original are close to their counterparts for simulated cascades,  $S$  will be near 1.0. In our case  $S = 1.0522$ , which means that most links, at least the most important links, have similar propagation capacity for the original and simulated cascades.

The CCD of the link weights for the original and simulated cascades are shown in Fig. 5(a). The two distributions match very well. Both of them follow obvious power law and can range from 1 to more than 1000, suggesting that a small number of links can cause much greater consequences than most of the others.

The CCD of the vertex out-strength and in-strength for original and simulated cascades are shown in Figs. 5(b)–5(c). Here the out-strength is defined as the summation of weights of the outgoing links from a vertex and the in-strength is the summation of weights of the incoming links to a vertex. An obvious power law behavior can be seen, which means that most vertices (component failures) have small consequences while a small number of them have much greater impact. Another point is that the strength distributions of the original and simulated cascades match very well, indicating that they share similar features from an overall point of view.

We eliminate 5% of the links (10 links) by setting 10 nonzero elements in  $\mathbf{B}$  matrix to be zero. We get  $\mathbf{B}_{\text{int}}$  by eliminating 10 links with the greatest weights and  $\mathbf{B}_{\text{rand}}$  by randomly removing 10 links. In Fig. 6 we show the position of the removed links in the interaction matrix. Then we separately simulate cascading failures with the proposed model by using  $\mathbf{B}_{\text{int}}$  and  $\mathbf{B}_{\text{rand}}$  and the two mitigation strategies are respectively called intentional mitigation and random mitigation. Fig. 7 shows the effects of the two mitigation strategies. It is seen that the risk of large-scale cascading failures can be significantly

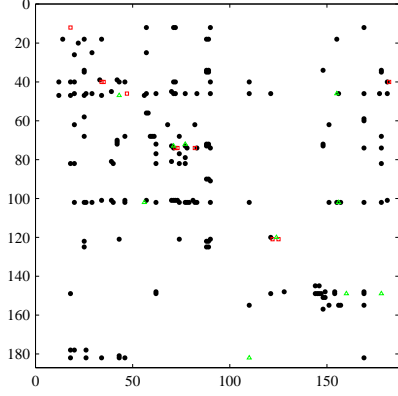


FIG. 6. Position of the removed links. Black dots denote nonzero elements of  $\mathbf{B}$ , among which the nonzero elements removed by intentional mitigation and random mitigation are denoted by red squares and green triangles.

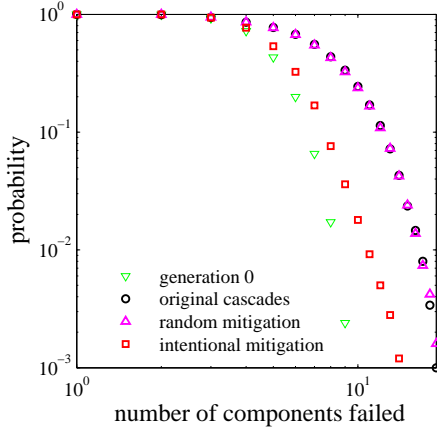


FIG. 7. CCD under two mitigation strategies.

mitigated by eliminating only a few most important links. By contrast, the mitigation effect is minor if we get rid of the same number of links randomly.

To conclude, we quantitatively determine the interaction of components in the system by calculating the probability that one component failure causes another. By using this information we propose a simple model to simulate cascading failures. The model is validated to be able to capture the general properties of the original cascades through comparison between the original and simulated cascades. An obvious power law is found in the distributions of the link weights and the vertex out-strength and in-strength. The links have significantly different propagation capacity. A small number of links are much more crucial than most of the others and by eliminating them the risk of cascading failures can be

dramatically mitigated.

We are grateful for the financial support of NSFC Grant No. 50525721, China's 863 Program Grant No. 2011AA05A118, and Henan cascading failure project.

- 
- [1] NERC (North America Electric Reliability Council), *1996 System Disturbances* (Princeton, 2002).
  - [2] U.S.-Canada Power System Outage Task Force, *Final Report on the August 14th Blackout in the United States and Canada* (2004).
  - [3] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, *Phys. Rev. Lett.* **87**, 258701 (2001).
  - [4] K. I. Goh, B. Kahng, and D. Kim, *Phys. Rev. Lett.* **88**, 108701 (2002).
  - [5] S. N. Dorogovtsev and J. F. Mendes, *Evolution of Networks* (Oxford University Press, Oxford, 2003).
  - [6] S. H. Strogatz, *Nature (London)* **410**, 268 (2001).
  - [7] R. N. Mantegna and H. E. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, Cambridge, 2000).
  - [8] M. Granovetter, *American J. Sociology*, **83**, 1420 (1978).
  - [9] S. Morris, *Review of Economic Studies*, **67**, 57 (2000).
  - [10] D. J. Watts, *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 5766 (2002).
  - [11] C. Asavathiratham, Ph.D. thesis, EECS Dept., MIT (2000).
  - [12] P. D. Hines, I. Dobson, E. Cotilla-Sanchez, and M. Eppstein, *Hawaii International Conference on System Sciences* (Maui, HI, 2013).
  - [13] C. Perrow, *Normal accident: living with high-risk technologies* (Princeton University Press, Princeton, 1999).
  - [14] I. Dobson, J. Kim, and K. R. Wierzbicki, *Risk Analysis*, **30**, 650 (2010).
  - [15] J. Kim, K. R. Wierzbicki, I. Dobson, and R. C. Hardiman, *IEEE Systems Journal*, **6**, 548 (2012).
  - [16] I. Dobson, *IEEE Trans. Power Syst.*, **27**, 2146 (2012).
  - [17] D. J. Watts and S. H. Strongatz, *Nature (London)* **393**, 440 (1998).
  - [18] R. Albert and A.-L. Barabási, *Rew. Mod. Phys.* **74**, 47 (2000).
  - [19] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 5766 (2004).
  - [20] A. Barrat, M. Barthélemy, and A. Vespignani, *Phys. Rev. Lett.* **92**, 228701 (2004).
  - [21] S. Mei, Y. Weng, X. Weng, and A. Xue, in *Proceedings of the 25th Chinese Control Conference*, 2006 (Harbin, China, 2006).
  - [22] S. Mei, Y. Ni, Weng, G. Wang, and S. Wu, *IEEE Trans. Power Syst.*, **23**, 1719 (2008).
  - [23] S. Mei, X. Zhang, and M. Cao, *Power Grid Complexity* (Tsinghua University Press, Beijing and Springer, 2011).
  - [24] H. Ren, I. Dobson, and B. A. Carreras, *IEEE Trans. Power Syst.*, **23**, 1217 (2008).
  - [25] B. A. Carreras, D. E. Newman, I. Dobson, and N. S. Degala, *46th Hawaii Intl. Conference on System Sciences*, 2013 (Maui, HI, 2013).
  - [26] The IEEE 118 Bus Test System. [Online]. Available: <http://www.ee.washington.edu/research/pstca/>.